

IN THE CLAIMS:

Please find below a listing of all pending claims. The statuses of the claims are set forth in parentheses. For those currently amended claims, underlined emphasis indicates insertions and ~~strikethrough~~ emphasis (and/or double brackets) indicates deletions.

1. (previously canceled)
2. (currently amended) [[A]]~~The text change frame detection apparatus that selects a plurality of video frames including text contents from given video frames, said apparatus according to claim 5, further comprising:~~
 - a first frame removing unit ~~removing to remove~~ the similar video frame as a redundant video ~~frames~~ frame from the given video frames;
 - a second frame removing unit ~~removing to remove~~ video frames that do not contain a text area from the given video frames; and
 - a third frame removing unit ~~detecting and removing to detect and remove~~ redundant video frames caused by image shifting from the given video frames[[; and]], wherein
 - [[an]]~~the output unit outputting to output~~ remaining video frames as the candidate text change frames[[,]]
~~wherein the first frame removing unit includes:~~
 - ~~an image block validation unit determining whether two image blocks in the same position in two video frames of the given video frames are a valid block pair that has an ability to show a change of image contents;~~
 - ~~an image block similarity measurement unit calculating a similarity of two image blocks of the valid block pair and determining whether the two image blocks are similar; and~~

a frame similarity judgment unit determining whether the two video frames are similar by using a ratio of a number of similar image blocks to a total number of valid block pairs;

and the first frame removing unit removes a similar video frame as a redundant video frame.

3. (currently amended) A text change frame detection apparatus that selects a plurality of video frames including text contents from given video frames, said apparatus comprising:

a first frame removing unit ~~removing to remove~~ redundant video frames from the given video frames;

a second frame removing unit ~~removing to remove~~ video frames that do not contain a text area from the given video frames;

a third frame removing unit ~~detecting and removing to detect and remove~~ redundant video frames caused by image shifting from the given video frames; and

an output unit ~~outputting to output~~ remaining video frames as candidate text change frames,

wherein the second frame removing unit includes:

a fast and simple image binarization unit ~~generating to generate~~ a first binary image of a video frame of the given video frames;

a text line region determination unit ~~determining to determine~~ a position of a text line region by using a horizontal projection and a vertical projection of the first binary image;

a rebinarization unit ~~generating to generate~~ a second binary image of every text line region;

a text line confirmation unit ~~determining to determine~~ validity of a text line region by using a difference between the first binary image and the second binary image and a fill rate of a number of foreground pixels in the text line region to a total number of pixels in the text line region; and

a text frame verification unit ~~confirming to confirm~~ whether a set of continuous video frames are non-text frames that do not contain a text area by using a number of valid text line regions in the set of continuous video frames.

4. (currently amended) A text change frame detection apparatus that selects a plurality of video frames including text contents from given video frames, said apparatus comprising:

a first frame removing unit ~~removing to remove~~ redundant video frames from the given video frames;

a second frame removing unit ~~removing to remove~~ video frames that do not contain a text area from the given video frames;

a third frame removing unit ~~deteeting and removing to detect and remove~~ redundant video frames caused by image shifting from the given video frames; and

an output unit ~~outputting to output~~ remaining video frames as candidate text change frames,

wherein the third frame removing unit includes:

a fast and simple image binarization unit ~~generating to generate~~ binary images of two video frames of the given video frames;

a text line vertical position determination unit ~~determining to determine~~ a vertical position of every text line region by using horizontal projections of the binary images of the two video frames;

a vertical shifting detection unit ~~determining to determine~~ a vertical offset of image shifting between the two video frames and a similarity of the two video frames in a vertical direction by using correlation between the horizontal projections; and

a horizontal shifting detection unit ~~determining to determine~~ a horizontal offset of the image shifting and a similarity of the two video frames in a horizontal direction by using correlation between vertical projections of every text line in the binary images of the two video frames,

and the third frame removing unit to remove[[s]] a similar video frame as a redundant video frame caused by the image shifting.

5. (currently amended) A text change frame detection apparatus that selects a plurality of video frames including text contents from given video frames, said apparatus comprising:

an image block validation unit determining whether to calculate a mean value and a variance of a gray level of each of two image blocks in the same position in two video frames of the given video frames, and to determine the two image blocks are a valid block pair that has an ability to show a change of image contents if at least one of two variances of the two image blocks is greater than a first threshold, or if the two variances are smaller than the first threshold and an absolute difference of two mean values of the two image blocks is greater than a second threshold;

an image block similarity measurement unit calculating to calculate a similarity of two image blocks of the valid block pair and to determine determining whether the two image blocks are similar;

a frame similarity judgment unit determining to determine whether the two video frames are similar by using a ratio of a number of similar image blocks to a total number of valid block pairs; and

an output unit outputting to output remaining video frames after a similar video frame is removed, as candidate text change frames.

6. (currently amended) A text change frame detection apparatus that selects a plurality of video frames including text contents from given video frames, said apparatus comprising:

a fast and simple image binarization unit generating to generate a first binary image of a video frame of the given video frames;

a text line region determination unit determining to determine a position of a text line region by using a horizontal projection and a vertical projection of the first binary image;

a rebinarization unit generating to generate a second binary image of every text line region;

a text line confirmation unit determining to determine validity of a text line region by using a difference between the first binary image and the second binary image and a fill rate of a number of foreground pixels in the text line region to a total number of pixels in the text line region;

a text frame verification unit confirming to confirm whether a set of continuous video frames are non-text frames that do not contain a text area by using a number of valid text line regions in the set of continuous video frames; and

an output unit outputting to output remaining video frames after the non-text frames are removed, as candidate text change frames.

7. (currently amended) A text change frame detection apparatus that selects a plurality of video frames including text contents from given video frames, said apparatus comprising:

a fast and simple image binarization unit generating to generate binary images of two video frames of the given video frames;

a text line vertical position determination unit determining to determine a vertical position of every text line region by using horizontal projections of the binary images of the two video frames;

a vertical shifting detection unit determining to determine a vertical offset of image shifting between the two video frames and a similarity of the two video frames in a vertical direction by using correlation between the horizontal projections;

a horizontal shifting detection unit determining to determine a horizontal offset of the image shifting and a similarity of the two video frames in a horizontal

direction by using correlation between vertical projections of every text line in the binary images of the two video frames; and

an output unit ~~outputting to output~~ remaining video frames after a similar video frame is removed, as candidate text change frames.

8. (withdrawn) A text extraction apparatus that extracts at least one text line region from a given image, said apparatus comprising:

an edge image generation unit generating edge information of the given image;

a stroke image generation unit generating a binary image of candidate character strokes in the given image by using the edge information;

a stroke filtering unit removing a false stroke from the binary image by using the edge information;

a text line region formation unit combining a plurality of strokes into a text line region;

a text line verification unit removing a false character stroke from the text line region and reforming the text line region;

a text line binarization unit binarizing the text line region by using a height of the text line region; and

an output unit outputting a binary image of the text line region.

9. (withdrawn) The text extraction apparatus according to claim 8, wherein the edge image generation unit includes:

an edge strength calculation unit calculating edge strength for every pixel in the given image by using a Sobel edge detector;

a first edge image generation unit generating a first edge image by comparing the edge strength of every pixel with a predefined edge threshold and setting a value of a corresponding pixel in the first edge image to one binary value if

the edge strength is greater than the threshold and the other binary value if the edge strength is less than the threshold; and

a second edge image generation unit generating a second edge image by comparing the edge strength of every pixel in a window centered at a position of every pixel of the one binary value in the first edge image with mean edge strength of the pixels in the window and setting a value of a corresponding pixel in the second edge image to the one binary value if the edge strength of the pixel is greater than the mean edge strength and the other binary value if the edge strength of the pixel is less than the mean edge strength.

10. (withdrawn) The text extraction apparatus according to claim 9, wherein the stroke image generation unit includes a local image binarization unit binarizing a gray scale image of the given image in a Niblack's binarization method to obtain the binary image of the candidate character strokes by using a window centered at a position of every pixel of the one binary value in the second edge image.

11. (withdrawn) The text extraction apparatus according to claim 9, wherein the stroke filtering unit includes:

a stroke edge coverage validation unit checking an overlap rate of a contour of a stroke in the binary image of the candidate character strokes by pixels of the one binary value in the second edge image, determining that the stroke is a valid stroke if the overlap rate is greater than a predefined threshold and an invalid stroke if the overlap rate is less than the predefined threshold, and removing the invalid stroke; and

a long straight line detection unit removing a large stroke by using a width and a height of the stroke.

12. (withdrawn) The text extraction apparatus according to claim 9, wherein the text line binarization unit includes:

an automatic size calculation unit determining a size of a window for binarization; and

a block image binarization unit binarizing a gray scale image of the given image in a Niblack's binarization method to obtain the binary image of the text line region by using the window centered at a position of every pixel of the one binary value in the second edge image.

13. (withdrawn) The text extraction apparatus according to claim 8, wherein the text line region formation unit includes a stroke connection checking unit checking whether two adjacent strokes are connectable by using an overlap ratio of heights of the two strokes and a distance between the two strokes, and the text line region formation unit combines the plurality of strokes into a text line region by using a result of checking.

14. (withdrawn) The text extraction apparatus according to claim 8, wherein the text line verification unit includes:

a vertical false stroke detection unit checking every stroke with a height higher than a mean height of strokes in the text line region, and marking the stroke as a false stroke if the stroke connects two horizontal text line regions into one big text line region;

a horizontal false stroke detection unit checking every stroke with a width larger than a threshold determined by a mean width of the strokes in the text line region, and marking the stroke as a false stroke if a number of strokes in a region that contains the stroke is less than a predefined threshold; and

a text line reformation unit reconnecting strokes except for a false stroke in the text line region if the false stroke is detected in the text line region.

15. (withdrawn) A text extraction apparatus that extracts at least one text line region from a given image, said apparatus comprising:

an edge image generation unit generating an edge image of the given image; a stroke image generation unit generating a binary image of candidate character strokes in the given image by using the edge image;

a stroke filtering unit checking an overlap rate of a contour of a stroke in the binary image of the candidate character strokes by pixels indicating an edge in the edge image, determining that the stroke is a valid stroke if the overlap rate is greater than a predefined threshold and an invalid stroke if the overlap rate is less than the predefined threshold, and removing the invalid stroke; and

an output unit outputting information of remaining strokes in the binary image of the candidate character strokes.

16. (previously canceled)

17. (currently amended) [[A]]~~The computer-readable storage medium storing a program used to direct a computer, that selects a plurality of video frames including text contents from given video frames, to perform a according to claim 20, wherein the process comprising further comprises:~~

~~removing the similar video frame as a redundant video frames frame~~ from the given video frames;

~~removing video frames that do not contain a text area from the given video frames; and~~

~~detecting and removing redundant video frames caused by image shifting from the given video frames[[; and]], wherein~~

~~the outputting outputs remaining video frames as the candidate text change frames[[,]]~~

~~wherein the removing redundant video frames includes:~~

~~determining whether two image blocks in the same position in two video frames of the given video frames are a valid block pair that has an ability to show a change of image contents;~~

~~calculating a similarity of two image blocks of the valid block pair and determining whether the two image blocks are similar; and~~

~~determining whether the two video frames are similar by using a ratio of a number of similar image blocks to a total number of valid block pairs, and the removing redundant video frames removes a similar video frame as a redundant video frame.~~

18. (previously presented) A computer-readable storage medium storing a program used to direct a computer, that selects a plurality of video frames including text contents from given video frames, to perform a process comprising:

removing redundant video frames from the given video frames;
removing video frames that do not contain a text area from the given video frames;

detecting and removing redundant video frames caused by image shifting from the given video frames; and

outputting remaining video frames as candidate text change frames, wherein the removing video frames that do not contain the text area includes:

generating a first binary image of a video frame of the given video frames;
determining a position of a text line region by using a horizontal projection and a vertical projection of the first binary image;

generating a second binary image of every text line region;

determining validity of a text line region by using a difference between the first binary image and the second binary image and a fill rate of a number of foreground pixels in the text line region to a total number of pixels in the text line region; and

confirming whether a set of continuous video frames are non-text frames that do not contain a text area by using a number of valid text line regions in the set of continuous video frames.

19. (previously presented) A computer-readable storage medium storing a program used to direct a computer, that selects a plurality of video frames including text contents from given video frames, to perform a process comprising:

removing redundant video frames from the given video frames;
removing video frames that do not contain a text area from the given video frames;
detecting and removing redundant video frames caused by image shifting from the given video frames; and

outputting remaining video frames as candidate text change frames, wherein the detecting and removing redundant video frames caused by image shifting includes:

generating binary images of two video frames of the given video frames;
determining a vertical position of every text line region by using horizontal projections of the binary images of the two video frames;
determining a vertical offset of image shifting between the two video frames and a similarity of the two video frames in a vertical direction by using correlation between the horizontal projections; and
determining a horizontal offset of the image shifting and a similarity of the two video frames in a horizontal direction by using correlation between vertical projections of every text line in the binary images of the two video frames, and the detecting and removing redundant video frames removes a similar video frame as a redundant video frame caused by the image shifting.

20. (currently amended) A computer-readable storage medium storing a program used to direct a computer, that selects a plurality of video frames including text contents from given video frames, to perform a process comprising:

~~determining whether calculating a mean value and a variance of a gray level of each of two image blocks in the same position in two video frames of the given video frames, and determining the two image blocks are a valid block pair that has~~

an ability to show a change of image contents if at least one of two variances of the two image blocks is greater than a first threshold, or if the two variances are smaller than the first threshold and an absolute difference of two mean values of the two image blocks is greater than a second threshold;

calculating a similarity of two image blocks of the valid block pair and determining whether the two image blocks are similar;

determining whether the two video frames are similar by using a ratio of a number of similar image blocks to a total number of valid block pairs; and

outputting remaining video frames after a similar video frame is removed, as candidate text change frames.

21. (original) A computer-readable storage medium storing a program used to direct a computer, that selects a plurality of video frames including text contents from given video frames, to perform a process comprising:

generating a first binary image of a video frame of the given video frames;

determining a position of a text line region by using a horizontal projection and a vertical projection of the first binary image;

generating a second binary image of every text line region;

determining validity of a text line region by using a difference between the first binary image and the second binary image and a fill rate of a number of foreground pixels in the text line region to a total number of pixels in the text line region;

confirming whether a set of continuous video frames are non-text frames that do not contain a text area by using a number of valid text line regions in the set of continuous video frames; and

outputting remaining video frames after the non-text frames are removed, as candidate text change frames.

22. (original) A computer-readable storage medium storing a program used to

direct a computer, that selects a plurality of video frames including text contents from given video frames, to perform a process comprising:

- generating binary images of two video frames of the given video frames;
- determining a vertical position of every text line region by using horizontal projections of the binary images of the two video frames;
- determining a vertical offset of image shifting between the two video frames and a similarity of the two video frames in a vertical direction by using correlation between the horizontal projections;
- determining a horizontal offset of the image shifting and a similarity of the two video frames in a horizontal direction by using correlation between vertical projections of every text line in the binary images of the two video frames; and
- outputting remaining video frames after a similar video frame is removed, as candidate text change frames.

23. (withdrawn) A computer-readable storage medium storing a program used to direct a computer, that extracts at least one text line region from a given image, to perform a process comprising:

- generating edge information of the given image; generating a binary image of candidate character strokes in the given image by using the edge information;
- removing a false stroke from the binary image by using the edge information;
- combining a plurality of strokes into a text line region;
- removing a false character stroke from the text line region and reforming the text line region;
- binarizing the text line region by using a height of the text line region; and
- outputting a binary image of the text line region.

24. (withdrawn) The storage medium according to claim 23, wherein the generating edge information includes:

calculating edge strength for every pixel in the given image by using a Sobel edge detector;

generating a first edge image by comparing the edge strength of every pixel with a predefined edge threshold and setting a value of a corresponding pixel in the first edge image to one binary value if the edge strength is greater than the threshold and the other binary value if the edge strength is less than the threshold; and

generating a second edge image by comparing the edge strength of every pixel in a window centered at a position of every pixel of the one binary value in the first edge image with mean edge strength of the pixels in the window and setting a value of a corresponding pixel in the second edge image to the one binary value if the edge strength of the pixel is greater than the mean edge strength and the other binary value if the edge strength of the pixel is less than the mean edge strength.

25. (withdrawn) The storage medium according to claim 24, wherein the generating the binary image of the candidate character strokes includes binarizing a gray scale image of the given image in a Niblack's binarization method to obtain the binary image of the candidate character strokes by using a window centered at a position of every pixel of the one binary value in the second edge image.

26. (withdrawn) The storage medium according to claim 24, wherein the removing the false stroke from the binary image includes:

removing a large stroke by using a width and a height of the stroke.
checking an overlap rate of a contour of a stroke in the binary image of the candidate character strokes by pixels of the one binary value in the second edge image;

determining that the stroke is a valid stroke if the overlap rate is greater than a predefined threshold and an invalid stroke if the overlap rate is less than the predefined threshold; and

removing the invalid stroke.

27. (withdrawn) The storage medium according to claim 24, wherein the binarizing the text line region includes:

determining a size of a window for binarization; and

binarizing a gray scale image of the given image in a Niblack's binarization method to obtain the binary image of the text line region by using the window centered at a position of every pixel of the one binary value in the second edge image.

28. (withdrawn) The storage medium according to claim 23, wherein the combining the plurality of strokes into the text line region includes checking whether two adjacent strokes are connectable by using an overlap ratio of heights of the two strokes and a distance between the two strokes, and the combining the plurality of strokes into the text line region combines the plurality of strokes into a text line region by using a result of checking.

29. (withdrawn) The storage medium according to claim 23, wherein the removing the false character stroke from the text line region and reforming the text line region includes:

checking every stroke with a height higher than a mean height of strokes in the text line region;

marking the stroke as a false stroke if the stroke connects two horizontal text line regions into one big text line region;

checking every stroke with a width larger than a threshold determined by a mean width of the strokes in the text line region;

marking the stroke as a false stroke if a number of strokes in a region that contains the stroke is less than a predefined threshold; and

reconnecting strokes except for a false stroke in the text line region if the false stroke is detected in the text line region.

30. (withdrawn) A computer-readable storage medium storing a program used to direct a computer, that extracts at least one text line region from a given image, to perform a process comprising:

- generating an edge image of the given image;
- generating a binary image of candidate character strokes in the given image by using the edge image;
- checking an overlap rate of a contour of a stroke in the binary image of the candidate character strokes by pixels indicating an edge in the edge image;
- determining that the stroke is a valid stroke if the overlap rate is greater than a predefined threshold and an invalid stroke if the overlap rate is less than the predefined threshold;
- removing the invalid stroke; and
- outputting information of remaining strokes in the binary image of the candidate character strokes.

31. (canceled)

32. (withdrawn) A text extraction method for extracting at least one text line region from a given image, said method comprising:

- generating edge information of the given image;
- generating a binary image of candidate character strokes in the given image by using the edge information;
- removing a false stroke from the binary image by using the edge information;
- combining a plurality of strokes into a text line region;
- removing a false character stroke from the text line region and reforming the text line region;

binarizing the text line region by using a height of the text line region; and
presenting a binary image of the text line region.